

Listing of Claims:

This listing of claims will replace all prior versions and listings of claims in the instant application:

1. (Previously Presented) A computer-implemented method for debugging and tuning synthesized audio, comprising the steps of:
 - (a) receiving a user-supplied text with a visual user interface;
 - (b) generating synthesized audio generated from concatenated phonetic units, the synthesized audio being a voice rendering of the user-supplied text;
 - (c) displaying a waveform corresponding to the synthesized audio generated from concatenated phonetic units;
 - (d) displaying parameters corresponding to at least one of the phonetic units, the parameters including configuration parameters comprising at least one weight for adjusting at least one search cost function, the at least one weight comprising at least one of a pitch cost weight and a duration cost weight;
 - (e) displaying an original recording containing a selected phonetic unit;
 - (f) receiving an editing input from the user;
 - (g) adjusting at least one configuration parameter in accordance with the editing input and storing the at least one configuration parameter in a text-to-speech engine configuration file, wherein adjusting includes repositioning a phonetic alignment marker;
 - (h) highlighting in the display of the original recording at least one user-selected phonetic unit;
 - (i) correcting elements of a text-to-speech segment dataset of parameters corresponding to a segment of the synthesized audio identified as be problematic;
 - (j) generating a new synthesized waveform corresponding to one or more adjusted parameters; and

(k) repeating steps (b)-(j) until a desired synthesized output is generated.

2. (Previously Presented) The method of claim 1, wherein said displaying parameters step further comprises displaying the parameters responsive to a user selection of at least a portion of the waveform, the displayed parameters correlating to the selected portion of the waveform.

3. (Original) The method of claim 1, wherein said displaying parameters step further comprises identifying a portion of the waveform responsive to a user selection of at least one of the parameters, the identified portion of the waveform correlating to the selected parameters.

4. (Cancelled)

5. (Cancelled)

6. (Cancelled)

7. (Cancelled)

8. (Previously Presented) The method of claim 1, wherein said adjusting step comprises at least one action selected from the group consisting of deleting a pitch mark, inserting a pitch mark, and repositioning a pitch mark by deleting a phonetic unit label, adding a phonetic unit label, modifying the phonetic unit label, and repositioning the phonetic unit boundaries.

9. (Previously Presented) The method of claim 1, wherein said displaying parameters step further comprises the step of displaying a waveform from the original recording along with the phonetic unit.
10. (Original) The method of claim 9, wherein edits to the waveform adjust parameters in the segment dataset.
11. (Cancelled)
12. (Cancelled)
13. (Previously Presented) A machine-readable storage having stored thereon a computer program having a plurality of code sections, the code sections executable by a machine for causing the machine to perform the steps of:
- (a) receiving a user-supplied text with a visual user interface;
 - (b) generating synthesized audio generated from concatenated phonetic units, the synthesized audio being a voice rendering of the user-supplied text;
 - (c) displaying a waveform corresponding to the synthesized audio generated from concatenated phonetic units;
 - (d) displaying parameters corresponding to at least one of the phonetic units, the parameters including configuration parameters comprising at least one weight for adjusting at least one search cost function, the at least one weight comprising at least one of a pitch cost weight and a duration cost weight;
 - (e) displaying an original recording containing a selected phonetic unit;
 - (f) receiving an editing input from the user;

(g) adjusting at least one configuration parameter in accordance with the editing input and storing the at least one configuration parameter in a text-to-speech engine configuration file, wherein adjusting includes repositioning a phonetic alignment marker;

(h) highlighting in the display of the original recording at least one user-selected phonetic unit;

(i) correcting elements of a text-to-speech segment dataset of parameters corresponding to a segment of the synthesized audio identified as be problematic;

(j) generating a new synthesized waveform corresponding to one or more adjusted parameters; and

(k) repeating steps (b)-(j) until a desired synthesized output is generated.

14. (Previously Presented) The machine-readable storage of claim 13, wherein said displaying parameters step further comprises displaying the parameters responsive to a user selection of at least a portion of the waveform, the displayed parameters correlating to the selected portion of the waveform.

15. (Original) The machine-readable storage of claim 13, wherein said displaying parameters step further comprises identifying a portion of the waveform responsive to a user selection of at least one of the parameters, the identified portion of the waveform correlating to the selected parameters.

16. (Cancelled)

17. (Cancelled)

18. (Cancelled)

19. (Cancelled)

20. (Previously Presented) The machine-readable storage of claim 13, wherein said adjusting step comprises at least one action selected from the group consisting of deleting a pitch mark, inserting a pitch mark, and repositioning a pitch mark by deleting a phonetic unit label, adding a phonetic unit label, modifying the phonetic unit label, and repositioning the phonetic unit boundaries.

21. (Previously Presented) The machine-readable storage of claim 13, wherein said displaying parameters step further comprises the step of displaying a waveform from the original recording along with the phonetic unit.

22. (Original) The machine-readable storage of claim 21, wherein edits to the waveform adjust parameters in the segment dataset.

23. (Cancelled)

24. (Cancelled)

25. (Previously Presented) A system for debugging and tuning synthesized audio, comprising:

means for receiving a user-supplied text with a visual user interface;

means for generating synthesized audio generated from concatenated phonetic units, the synthesized audio being a voice rendering of the user-supplied text;

means for displaying the waveform corresponding to synthesized audio generated from concatenated phonetic units;

means for displaying parameters corresponding to at least one of the phonetic units, the parameters including configuration parameters comprising at least one weight for adjusting at least one search cost function, the at least one weight comprising at least one of a pitch cost weight and a duration cost weight;

means for displaying an original recording containing a selected phonetic unit;

means for receiving an editing input from the user; and means for adjusting the parameters in accordance with the editing input by adjusting and storing in a text-to-speech engine configuration file at least one configuration parameter, wherein adjusting includes repositioning a phonetic alignment marker.

means for highlighting in the display of the original recording at least one user-selected phonetic unit;

means for correcting elements of a text-to-speech segment dataset of parameters corresponding to a segment of the synthesized audio identified as be problematic;

means for generating a new synthesized waveform corresponding to one or more adjusted parameters; and

wherein the system continues to regenerate new synthesized waveforms until a desired synthesized output is generated.

26. (Previously Presented) The method of Claim 1 wherein the parameter updates and segment dataset corrections are applied in regenerating the synthesized audio.